EVOLUTION OF SOCIAL P2P NETWORKS BASED ON THE DYNAMICS OF HETEROGENEOUS MULTIMEDIA PEERS

Hyunggon Park and Mihaela van der Schaar

Multimedia Communications and Systems Lab.
Electrical Engineering Department, University of California, Los Angeles (UCLA)

ABSTRACT

In this paper, we consider social peer-to-peer (P2P) networks, where peers are sharing their resources (i.e., multimedia content and upload bandwidth). In the considered P2P networks, peers are self-interested, thereby determining their resource divisions (i.e., actions) among their associated peers such that their utility (e.g., multimedia quality) is maximized. Peers determine their optimal strategies for selecting their action based on a Markov Decision Process (MDP) framework, which enables the peers to maximize their cumulative utilities. We consider heterogeneous peers that have different and limited ability to characterize their resource reciprocations using only a limited number of states. We investigate how the limited number of states impacts the resource reciprocation and the resulting multimedia quality over time. Simulation results show that peers simultaneously refining their state descriptions can improve the multimedia quality in the resource reciprocation. Moreover, peers prefer to interact with other peers that have higher available upload bandwidths as well as have similar capabilities for refining their number of states.

Index Terms— Social peer-to-peer (P2P) network, evolution of resource reciprocation, Markov decision process.

1. INTRODUCTION

Social network communities such as [1–3] have recently become popular, and among them, peer-to-peer (P2P) applications represent a large majority of the traffic currently transmitted over the Internet. The traffic exchanged is often multimedia content, e.g. downloads of multimedia data.

Recently, several solutions have been proposed for general file sharing [2,4,5] and multimedia streaming [4,6] in P2P networks. Among these solutions, we consider data-driven approaches [4–6], where multimedia content or general files of each peer are divided into chunks of uniform length and are then disseminated over the P2P network. Based on the chunk availability, peers form groups with which they can continuously exchange their chunks. While this approach has been successfully deployed in P2P applications, key challenges such as determining optimal resource reciprocation strategies among self-interested peers still remain largely unaddressed.

A resource reciprocation strategy among self-interested peers in BitTorrent systems has been developed based on a tit-for-tat (TFT) strategy, where a peer selects some of its associated peers (i.e., leechers) which are *currently* uploading at the highest rates to download its content [5]. A key disadvantage of this method is that a peer deploying this strategy decides its resource reciprocation by evaluating only the current upload rates which it receives from its associated peers. Thus, the resource reciprocation is determined *myopically*. However, since peers in P2P networks are generally involved in repeated, long-term interactions, such myopic decisions can result in a suboptimal performance for the involved peers.

To take into account the repeated resource reciprocation among self-interested peers, each peer determines its actions by considering the probabilistic behavior of resource reciprocation of its associated peers. Formalizing the resource reciprocation based on a MDP [7] enables the peers to take their foresighted actions in a way that maximizes their expected cumulative rewards (e.g., download rates or multimedia quality). While our previous work [8] shows that the MDP-based foresighted strategies improve the performance of the P2P applications, it does not investigate how heterogeneous peers interact with each other based on their different abilities.

We consider heterogeneous peers that have different and limited abilities to characterize their resource reciprocation. The resource reciprocation of each peer is described based on a finite number of *state descriptions*. Hence, the heterogeneous peers cannot differentiate among all possible download rates from their associated peers. Consequently, a peer may have multiple actions that are optimal because these actions do not alter its associated peers' states, and thus, they do not alter the resource reciprocation of these peers. We analytically show that multiple optimal actions exist for such heterogeneous peers. Moreover, we show that peers can mutually improve their download rates only if they simultaneously refine their state descriptions.

This paper is organized as follows. In Section 2, the MDP-based resource reciprocation strategy for P2P networks is presented. In Section 3, we study the evolution of resource reciprocation for heterogeneous peers. Simulation results are presented in Section 4 and conclusions are drawn in Section 5.

2. RESOURCE RECIPROCATION STRATEGY BASED ON MDP

2.1. Resource Reciprocation Model

In data-driven P2P networks, the set of peers which share their content and resources (i.e., associated peers) can be considered as a *group*[1]. The group of peer i that consists of N_{C_i} peers is denoted by $C_i = \{1, \ldots, N_{C_i}\}$. We assume that each peer decides its resource reciprocation in its group using an MDP, which is explained in the following. A similar MDP-based resource reciprocation model can be found in our previous work [8].

For a peer i, an MDP is a tuple $\langle \mathbf{S}_i, \mathbf{A}_i, P_i, R_i \rangle$, where $\mathbf{S}_i$ is the state space, $\mathbf{A}_i$ is the action space, $P_i : \mathbf{S}_i \times \mathbf{A}_i \times \mathbf{S}_i \to [0, 1]$ is a state transition probability function, and $R_i : \mathbf{S}_i \to \mathbb{R}^+$ is the reward derived in state $s_i \in \mathbf{S}_i$.

· *State Space* $\mathbf{S}_i$: A state of peer i represents the received resources from the peers in C_i. Peer i's resources received from its associated peers in C_i can be expressed as $(x_{1i}, \ldots, x_{N_{C_i} i})$, where x_{ki} ($0 \leq x_{ki} \leq L_k$) denotes the resources provided by peer $k \in C_i$ to peer i, and L_k represents the available maximum upload bandwidth of peer k. For peer i, its download rates are recognized and represented by n_i discrete values based on its quantization function ψ_i, i.e., a download rate from peer x_{ki} is mapped into $\psi_i(x_{ki}) = s_{ik} \in \{s_{ik}^1, \ldots, s_{ik}^{n_i}\}$. Hence, the state space of peer i can be expressed as

$$\mathbf{S}_i = \left\{ s_i = (s_{i1}, \ldots, s_{iN_{C_i}}) \middle| s_{ik} = \psi_i(x_{ki}), k \in C_i \right\}. \quad (1)$$

The set of n_i discrete values of $\{s_{ik}^1, \ldots, s_{ik}^{n_i}\}$ is referred to as *state descriptions* in this paper.

· *Action Space* $\mathbf{A}_i$: An action of peer i is its resource allocation to the associated peers in C_i. A set of resource allocations of peer i in C_i can be expressed as $(x_{i1}, \ldots, x_{iN_{C_i}})$, where $\sum_{k \in C_i} x_{ik} \leq L_i$. We assume that the download bandwidth of each peer is larger than the upload bandwidth, and the upload bandwidth is decomposed into "units" of bandwidth, denoted by Δx_i. Thus, action $\mathbf{a}_i = (a_{i1}, \ldots, a_{iN_{C_i}})$, where $a_{il} \in \mathbb{Z}^+$, is the number of units of bandwidth that is allocated to the associated peers. Thus,

$$\mathbf{A}_i = \{\mathbf{a}_i | a_{ik} \leq a_i^M, \sum_{k \in C_i} a_{ik} \leq a_i^M, k \in C_i\}, \quad (2)$$

where $a_i^M = \lfloor L_i/\Delta x_i \rfloor$. Hence, action a_{ik} determines peer k's download rate from peer i, i.e., $x_{ik} = a_{ik}\Delta x_i$.

· *State Transition Probability* $P_{\mathbf{a}_i}(s_i, s_i')$: A state transition probability represents the probability that an action of a peer in a state will lead to another state. In the MDP model, given a state $s_i \in \mathbf{S}_i$ at time t, an action $\mathbf{a}_i \in \mathbf{A}_i$ of peer i can lead to another state $s_i' \in \mathbf{S}_i$ at $t + 1$ with probability $P_{\mathbf{a}_i}(s_i, s_i')$, defined as

$$P_{\mathbf{a}_i}(s_i, s_i') = \Pr(s_i^{(t+1)} = s_i' | s_i^{(t)} = s_i, \mathbf{a}_i^{(t+1)} = \mathbf{a}_i). \quad (3)$$

In this paper, we assume that the state transition probabilities of peers are known; that is, they have been efficiently identified based on e.g., [8], and available.

· *Reward* R_i: The reward of a peer in a state represents its total download rate in that state. Since the state s_i of peer i is determined by the download rates from C_i using function ψ_i, the download rates in a state can be explicitly determined. Hence, the reward in a state $s_i = (s_{i1}, \ldots, s_{iN_{C_i}})$ can be expressed as

$$R(s_i) = \sum_{k \in C_i} r_i(s_{ik}), \quad (4)$$

where $r_i(s_{ik})$ is a random variable that represents download rates from peer k in s_{ik}.

· *Resource Reciprocation Policy* π_i^*: The solution to the MDP is represented by peer i's optimal resource reciprocation policy π_i^*, which maps the states to optimal actions. The optimal policy can be obtained using well-known methods such as value iteration or policy iteration [7]. Hence, peer i can decide its actions based on the optimal policy π_i^*, i.e.,

$$\pi_i^*(s_i) = \mathbf{a}_i, \quad (5)$$

for all $s_i \in \mathbf{S}_i$. We assume that the optimal policy is continuously updated to capture the changes in the resource reciprocation.

2.2. Optimal Resource Reciprocation Strategy

Peers that use the MDP framework for their resource reciprocation aim to maximize their cumulative (discounted) expected download rates (CEDRs) by strategically allocating their upload bandwidth. The CEDRs at time t_c can be expressed as

$$R_i^{CEDR}(s_i^{(t_c)}) = \sum_{t=t_c+1}^{\infty} \gamma_i^{(t-(t_c+1))} \cdot E[R(s_i^{(t)})], \quad (6)$$

where constant γ_i ($0 \leq \gamma_i < 1$) is referred to as a *discount factor*, which determines a tradeoff between immediate reward and future reward[2], and $R(s_i^{(t)}) = \sum_{l=1}^{N_{C_i}} r_i(s_{il})$ for $s_i^{(t)} = (s_{i1}, \ldots, s_{iN_{C_i}})$. Hence, peer i can determine a set of actions that maximizes $R_i^{CEDR}(s_i^{(t)})$ in (6) for every state $s_i \in \mathbf{S}_i$, which leads to an optimal policy π_i^*. The optimal policy π_i^* thus maps each state $s_i \in \mathbf{S}_i$ into a corresponding optimal action $\mathbf{a}_i^*$, i.e., $\pi_i^*(s_i) = \mathbf{a}_i^*$ given a number of state descriptions.

It can be observed from (6) that the granularity of states (i.e., the number of state descriptions) of the peers impacts their policy as well as their corresponding actions, which ultimately affects the resulting download rates of all peers. This is because an optimal policy is mainly affected by the state transition probability. In the next section, we study the impact of the degree of granularity of state descriptions and its impact on evolution of resource reciprocation.

---

[1] In this paper, a set of associated peers and a group are interchangeably used. Groups in P2P networks can be swarms [4] or partnerships [6].

[2] The discount factor γ can represent validity of the expected future rewards. For example, a small discount factor is desirable for a transient regime, while a large discount factor can be used for a stationary regime [9].

3. EVOLUTION OF RESOURCE RECIPROCATION

As discussed in Section 2, given a number of state descriptions, each peer can determine its optimal strategy and the corresponding optimal actions based on the proposed MDP-based approach. In this section, we show that the peers can additionally improve their mutual download rates by evolving their resource reciprocation strategies if multiple optimal actions are available to them. Since the complexity for finding the optimal policy depends on the number of states[3], the maximum available number of state descriptions may not be feasible for each peer. Note that the peers do not need to compute their new policies while evolving their resource reciprocation.

3.1. Different State Granularity for Heterogeneous Peers

As discussed in Section 2.1, for a peer i with n_i state descriptions, a function ψ_i can map a download rate x_{ki} from peer k into a state description s_{ik}, i.e., $\psi_i(x_{ki}) = s_{ik} \in \{s_{ik}^1, \ldots, s_{ik}^{n_i}\}$. Specifically, we use a uniform quantization function ψ_i for maximum and minimum desired download rates of peer i, L_i^M and L_i^m, defined as

$$\psi_i(x_{ki}) \tag{7}$$
$$= \begin{cases} s_{ik}^1, \text{if } x_{ki} \in [0, L_i^m + \frac{1}{n_i}\overline{L_k^m L_k^M}], \\ s_{ik}^l, \text{if } x_{ki} \in [L_i^m + \frac{l-1}{n_i}\overline{L_k^m L_k^M}, L_i^m + \frac{l}{n_i}\overline{L_k^m L_k^M}], \\ s_{ik}^{n_i}, \text{if } x_{ki} \in [L_i^m + \frac{l}{n_i}\overline{L_k^m L_k^M}, L_i], \end{cases}$$

for $2 \leq l \leq n_i - 1$ and $\overline{L_k^m L_k^M} = L_k^M - L_k^m$. As shown in (7), if a smaller number of state descriptions is used given L_i^M and L_i^m, each state description represents a larger interval, which leads to peer i becoming more insensitive to download rate variations from its associated peers. Hence, if peers with different number of state descriptions are reciprocating their resources, the peers having more state descriptions can have multiple actions that are optimal to their associated peers.

3.2. Multiple Optimal Actions

Let us consider a peer i with the unit of upload bandwidth Δx_i in a state $s_i \in \mathbf{S}_i$ interacting with the peers $k \in C_i$. An optimal action $\mathbf{a}_i^*$ is determined by its optimal policy π_i^*.

Proposition 1. *If $\Delta x_i \leq \overline{L_k^m L_k^M}/2n_k$, then there exists multiple optimal actions for peer i in each of states.*

Proof. Let $\mathbf{a}_i^* = (a_{i1}^*, \ldots, a_{iN_{C_i}}^*)$ be an optimal action of a peer i determined by π_i^*, i.e., $\pi_i^*(s_i) = \mathbf{a}_i^*$. Suppose that a_{ik}^* is mapped into s_{ki}^l by peer k, i.e., $\psi_k(a_{ik}^* \Delta x_i) = s_{ki}^l$. Then, $a_{ik}^* \Delta x_i \in [L_k^m + \frac{l-1}{n_k}\overline{L_k^m L_k^M}, L_k^m + \frac{l}{n_k}\overline{L_k^m L_k^M}] \triangleq \mathcal{R}_k^l$. Since $\Delta x_i \leq \overline{L_k^m L_k^M}/2n_k$,

$$(a_{ik}^* + 1)\Delta x_i \in \mathcal{R}_k^l,$$

---
[3]For the value iteration algorithm, the complexity is quadratic in the number of states [10].

if $a_{ik}^* \Delta x_i \in [L_k^m + \frac{l-1}{n_k}\overline{L_k^m L_k^M}, L_k^m + \frac{2l-1}{2n_k}\overline{L_k^m L_k^M}]$, i.e., $\psi_k(a_{ik}^* + 1) = \psi_k(a_{ik}^*)$. Alternately,

$$(a_{ik}^* - 1)\Delta x_i \in \mathcal{R}_k^l,$$

if $a_{ik}^* \Delta x_i \in [L_k^m + \frac{2l-1}{2n_k}\overline{L_k^m L_k^M}, L_k^m + \frac{l}{n_k}\overline{L_k^m L_k^M}]$, i.e., $\psi_k(a_{ik}^* - 1) = \psi_k(a_{ik}^*)$. Therefore, if $\Delta x_i \leq \overline{L_k^m L_k^M}/2n_k$, there exist actions $\mathbf{a}_i = (a_{i1}, \ldots, a_{iN_{C_i}}) \neq \mathbf{a}_i^*$ that cannot be differentiated by peers $k \in C_i$. Hence, the actions $\mathbf{a}_i$ are also optimal, concluding that peer i can have multiple optimal actions given its state $s_i \in \mathbf{S}_i$ and optimal policy π_i^*. □

Proposition 1 shows that peers can have multiple optimal actions, which enable them to maximize their CEDRs in their groups. Hence, given a policy π_i^*, using different optimal actions results in the same CEDRs for peer i, unless the associated peers refine their state descriptions such that they can differentiate the actions.

If several peers can simultaneously refine their state descriptions, and thus, they can differentiate their associated peers' optimal actions, then they can select actions beneficial to each other. This leads them to mutually improve their download rates, i.e., the resource reciprocation evolves. On the other hand, peers that use a fixed number of state descriptions in their resource reciprocation will be penalized by their associated peers that can keep refining their state descriptions. The evolution of resource reciprocation is quantitatively evaluated through simulation results in Section 4.

4. SIMULATION RESULTS

To highlight the impact of the evolution of resource reciprocation, we assume that each peer currently adopts the optimal policy and the initial resource reciprocation is determined by the policy. The policy is implemented based on the value iteration method [7].

4.1. Evolution of Resource Reciprocation and its Impact on Multimedia Quality

In this simulation, we consider an illustrative group with three peers, which have multiple optimal actions and can refine their state descriptions over time. We assume that $L_i^m = 0$ and $L_i^M = 208$Kbps, 222Kbps, and 258Kbps for peer $i = 1, 2, 3$, respectively. The unit of bandwidth is 2Kbps, ($\Delta x_i = 2$Kbps), and initial numbers of state descriptions are 2, 2, and 2 for peer 1, 2, and 3, respectively. For illustration, in the considered group, peer 1 and peer 2 can refine their state descriptions, while peer 3 cannot. Without loss of generality, we assume that the peers simultaneously refine their state description by doubling the number of state descriptions for each state (i.e., using 1 bit more for state descriptions) at their 15, 50, 75, and 95th resource reciprocation. We assume that a video sequence (*Foreman*) at CIF (352×288) resolution, 30
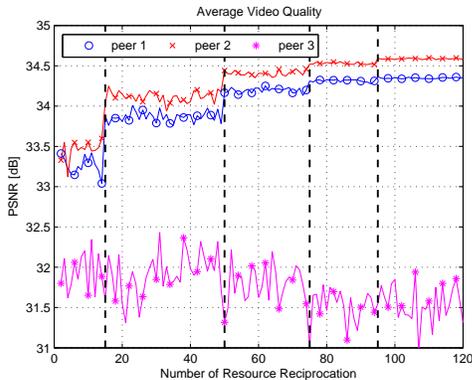
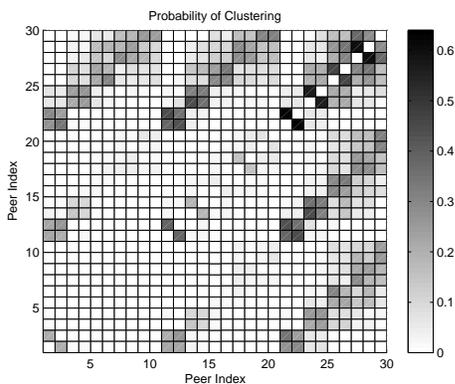**Fig. 1**. Average video quality (*Foreman sequence*).

Fig. 2. Probability of clustering among heterogeneous peers.

frames/sec, and encoded in a prioritized manner using the H.264/AVC encoder is exchanged. The encoded video file is partitioned into chunks that have uniform size of 20Kbits.

Simulation results are shown in Fig. 1, which shows the resulting multimedia quality measured in PSNR (Peak Signal to Noise Rati) among the peers in the group. Fig. 1 clearly shows that two peers (i.e., peer 1 and peer 2) can improve their multimedia qualities, by simultaneously increasing the number of state descriptions, while peer 3 cannot. Hence, the average download rates of peer 1 and peer 2 improve as refined state descriptions are used. Note that we can observe that the improvement by increasing the number of state descriptions in this resource reciprocation process is bounded, as this resource reciprocations are initiated by a fixed number of state descriptions that is determined by the MDP.

4.2. Clustering among Heterogeneous Peers

In this section, we study how heterogeneous peers are clustered, forming their groups. While forming groups, they consider the expected download rates achieved by the CEDR as well as the evolution of resource reciprocation.

Fig. 2 shows the distribution of probabilities that each peer forms a group with the other peers. To highlight the impact

of peers' heterogeneity on clustering, we assume that peers are classified into three categories based on their ability to refine their state descriptions or based on their available upload bandwidth. Peers in the same class have the same maximum available bandwidth. However, peers in a category are divided into 5 sub-classes based on their ability to refine their state descriptions. Two peers are included in each sub-class, and the peers with higher peer index in each category can refine more state descriptions. From the results shown in Fig. 2, we can also verify that peers prefer to make groups with peers who have the same ability to refine their state descriptions, and who have a higher available bandwidth.

5. CONCLUSIONS

We study the interactions among self-interested peers by sharing their resources in P2P networks. We show that MDP-based resource reciprocation strategy for peers' upload bandwidth allocation enables them to maximize their long-term multimedia quality. We analytically investigate the impact of peers' heterogeneous ability to refine their state descriptions on their interactions of resource reciprocation. We show that peers can achieve higher multimedia quality if they simultaneously improve their state descriptions. Finally, group formation for heterogeneous peers is discussed, showing that peers prefer to form groups with other peers that have a similar ability to refine state descriptions, but also have higher upload bandwidths.

6. REFERENCES

[1] "Napster." [Online]. Available: http://www.napster.com

[2] "KaZaA." [Online]. Available: http://www.kazaa.com

[3] "YouTube." [Online]. Available: http://www.youtube.com

[4] A. Legout, N. Liogkas, E. Kohler, and L. Zhang, "Clustering and sharing incentives in BitTorrent systems," *SIGMETRICS Perform. Eval. Rev.*, vol. 35, no. 1, pp. 301–312, 2007.

[5] B. Cohen, "Incentives build robustnessin BitTorrent," in *Proc. P2P Economics Workshop*, Berkerly, CA, 2003.

[6] X. Zhang, J. Liu, B. Li, and T. S. P. Yum, "CoolStreaming/DONet: A data-driven overlay network for efficient live media streaming," in *Proc. IEEE INFOCOM '05*, vol. 3, Mar. 2005, pp. 2102–2111.

[7] D. P. Bertsekas, *Dynamic Programming and Stochastic Control.* Academic Press, 1976.

[8] H. Park and M. van der Schaar, "Foresighted resource reciprocation strategies in P2P networks," in *Proc. IEEE Globecom '08*, Dec. 2008.

[9] G. de Veciana and X. Yang, "Fairness, incentives and performance in peer-to-peer networks," in *41th Ann. Allerton Conf. on Commun. , Control and Comput.*, 2003.

[10] M. L. Littman, T. L. Dean, and L. P. Kaelbling, "On the complexity of solving Markov decision problems," in *Proc. 11th Conf. on Uncertainty in Artificial Intelligence*, May 1995.